

CLAIMS

What is claimed is:

1. A media production system, comprising:
 - a textual alignment module aligning multiple speech recordings to textual lines of a script based on speech recognition results;
 - a navigation module responding to user navigation selections respective of the textual lines of the script by communicating to the user corresponding, line-specific portions of the multiple speech recordings; and
 - an editing module responding to user associations of multiple speech recordings with textual lines by accumulating line-specific portions of the multiple speech recordings in a combination recording based on at least one of relationships of textual lines in the script to the combination recording, and temporal alignments between the multiple speech recordings and the combination recording.
2. The system of claim 1, further comprising a ranking module adapted to tag at least one of speech recordings and specific portions thereof with ranking data.
3. The system of claim 2, wherein said ranking module is adapted to recognize tags associated with the speech recordings and tag at least one of speech recordings and specific portions thereof accordingly.

4. The system of claim 3, wherein said ranking module is adapted to recognize voice tags based on key phrases.

5. The system of claim 2, wherein said ranking module is adapted to recognize key phrases within the speech recordings and tag at least one of speech recordings and specific portions thereof accordingly.

6. The system of claim 2, wherein said ranking module is adapted to evaluate pitch of speech within the speech recordings and tag at least one of speech recordings and specific portions thereof accordingly.

7. The system of claim 2, wherein said ranking module is adapted to evaluate speed of speech within the speech recordings and tag at least one of speech recordings and specific portions thereof accordingly.

8. The system of claim 2, wherein said ranking module is adapted to evaluate emotive character of speech within the speech recordings and tag at least one of speech recordings and specific portions thereof accordingly.

9. The system of claim 1, wherein said navigation module is adapted to rank at least one of speech recordings and specific portions thereof based on predetermined ranking criteria and at least one of:

- (a) characteristics of at least one of speech recordings and specific portions thereof; and
- (b) ranking data associated with at least one of speech recordings and specific portions thereof.

10. The system of claim 9, further wherein said navigation module is adapted to rank at least one of speech recordings and specific portions thereof based on order in which the speech recordings were produced.

11. The system of claim 9, wherein said navigation module is adapted to rank at least one of speech recordings and specific portions thereof based on quality of pronunciation of speech therein.

12. The system of claim 9, wherein said navigation module is adapted to rank at least one of speech recordings and specific portions thereof based on pitch of speech therein.

13. The system of claim 9, wherein said navigation module is adapted to rank at least one of speech recordings and specific portions thereof based on speed of speech therein.

14. The system of claim 9, wherein said navigation module is adapted to rank at least one of speech recordings and specific portions thereof based on duration thereof.

15. The system of claim 9, wherein said navigation module is adapted to rank a line-specific portion of a speech recording based on consistency thereof with at least one adjacent, line-specific portion of another speech recording already assigned to a textual line sequentially adjacent in the script to a textual line aligned to the line-specific portion of the speech recording.

16. The system of claim 9, wherein said navigation module is adapted to rank at least one of speech recordings and specific portions thereof based on ability of thereof to contribute to solutions rendering a combination recording of a target duration and including a partial accumulation of line-specific portions of the multiple speech recordings.

17. The system of claim 9, wherein said navigation module is adapted to rank at least one of speech recordings and specific portions thereof based on ranking tags supplied thereto by speech recording production personnel during a speech recording process.

18. The system of claim 9, wherein said navigation module is adapted to rank at least one of speech recordings and specific portions thereof based on emotive character exhibited thereby and a target emotive state recorded with respect to a textual line aligned thereto.

19. The system of claim 9, wherein said navigation module is adapted to rank at least one of speech recordings and specific portions thereof in accordance with user-specified weights respective of multiple ranking criteria.

20. The system of claim 9, wherein said navigation module is adapted to automatically select at least one of speech recordings and specific portions thereof based on the predetermined ranking criteria.

21. The system of claim 1, wherein said navigation module is adapted to play a user-specified portion of a speech recording in response to a sample request.

22. The system of claim 1, wherein said navigation module is adapted to play at least one of a user-specified section of the combination recording and a preview of the user-specified section based on a sequence of portions of multiple speech recordings.

23. The system of claim 1, wherein said navigation module is adapted to record final selection of at least one of a speech recording and a specific portion thereof with respect to a textual line.

24. The system of claim 1, wherein the combination recording includes at least one voice track of a multiple track audio visual recording, the speech recordings are produced in a dubbing process, and each speech recording is automatically temporally aligned to the combination recording during the dubbing process.

25. The system of claim 1, wherein the textual lines are sequentially related and the combination recording includes at least one audio track having a durational constraint.

26. The system of claim 1, wherein the combination recording includes a navigable set of voice prompts.

27. The system of claim 1, wherein the combination recording includes a set of training data for at least one of a speech synthesizer and a speech recognizer.